Full Title: Post-Extinction Conditional Stimulus Valence Predicts Reinstatement Fear: Relevance for Long Term Outcomes of Exposure Therapy

Tomislav D. Zbozinek, B.A.[1]
Dirk Hermans, Ph.D.[2]
Jason M. Prenoveau, Ph.D.[3]
Betty Liao, M.A.[1]
Michelle G. Craske, Ph.D.[1]

1. University of California, Los Angeles, Department of Psychology
2. University of Leuven, Department of Psychology and Educational Sciences
3. Loyola University Maryland, Department of Psychology

Correspondence to:
Michelle G. Craske, Ph.D.
Professor
Department of Psychology
University of California, Los Angeles
1285 Franz Hall
Los Angeles, California 90095
Phone: 310-825-8403
Email: craske@psych.ucla.edu

Acknowledgements

The research reported and the preparation of this article was supported by National Institute of Mental Health Grants R21 MH722 59-01 to Michelle G. Craske and Mark Barad, by support from the Virginia Friedhofer Charitable Trust to Edward M. Ornitz, and by support from the National Science Foundation to Jason M. Prenoveau.

Abstract

Exposure therapy for anxiety disorders is translated from fear conditioning and extinction. While exposure therapy is effective in treating anxiety, fear sometimes returns after exposure. One pathway for return of fear is reinstatement: unsignaled unconditional stimuli following completion of extinction. The present study investigated the extent to which valence of the conditional stimulus (CS+) after extinction predicts return of CS+ fear after reinstatement. Participants (N = 84) engaged in a differential fear conditioning paradigm and were randomized to reinstatement or non-reinstatement. We hypothesized that more negative post-extinction CS+ valence would predict higher CS+ fear after reinstatement relative to non-reinstatement and relative to extinction retest. Results supported the hypotheses and suggest that strategies designed to decrease negative valence of the CS+ may reduce the return of fear via reinstatement following exposure therapy.

Keywords: fear conditioning, exposure therapy, anxiety, reinstatement, extinction, return of fear

Introduction

Exposure therapy is well established as an effective therapeutic strategy for anxiety disorders (Hofmann & Smits, 2008; In-Albon & Schneider, 2007). However, a certain proportion of individuals experience a return of fear following successful conclusion of treatment (Craske & Mystkowski, 2006; Rachman, 1989). Thus, there is a need to understand the mechanisms responsible for return of fear and to develop interventions that may reduce its occurrence. Extinction-based models of exposure therapy provide mechanisms that explain return of fear. The goal of the current study is to expand upon existing experimental findings suggesting that increased negative valence of a previously feared stimulus may contribute to return of fear following an unpredicted aversive event that occurs after completion of exposure therapy.

Exposure therapy is translated from fear conditioning and extinction. It is now thought that inhibitory learning is central to extinction (Bouton, 1993; Wagner, 1981), although additional mechanisms, such as habituation, may also be involved (Myers & Davis, 2007). Within a classical conditioning approach, the inhibitory learning models mean that the original conditional stimulus (CS) / unconditional stimulus (US) association learned during fear conditioning is not erased during extinction, but rather is left intact while a new, secondary CS/No-US inhibitory association develops (e.g., Bouton, 1993; Bouton & King, 1983). The inhibitory association is dependent on both the CS and the context in which the CS is presented, whereas the initial excitatory association is independent of context (Bouton, 2004). Bouton and colleagues propose that after extinction, the CS possesses two meanings: its original excitatory meaning (CS/US) and an additional inhibitory meaning (CS/No-US). Therefore, even though fear subsides with enough extinction trials, retention of at least part of the CS/US association can be uncovered after extinction.

One way that conditional fear can return after extinction is spontaneous recovery (Quirk, 2002), meaning that the strength of the conditional response (CR; i.e., fear) increases in proportion to the amount of time that passes between extinction and extinction retest, which is assessed at least 24 hours after extinction training. Clinically, this effect parallels the return of fear that commonly occurs with lengthy intervals of time since the end of exposure therapy and the first time a previously feared CS is re-encountered. Second, return of fear via renewal may occur due to a change in context between extinction and extinction retest (Bouton, 1993). Using a clinical example, an individual who undergoes exposure to public speaking in a therapy setting may experience a renewal of fear when speaking at a wedding (e.g., Culver, et al., 2011). Finally, return of fear may occur as a result of reinstatement of conditional fear due to unsignaled (or, unpaired) US presentations after extinction (Rescorla & Heth, 1975). For example, an individual who is afraid of the physical pain/bodily harm (US) associated with being bitten by a snake (CS+) may experience reinstatement of fear of the snake if they experience physical pain/bodily harm from a car accident. Reinstatement has been long established in animal fear conditioning studies and more recently has been shown in human conditioning studies (e.g., Dirikx et al., 2004; Dirikx et al., 2007; Hermans et al., 2005; LaBar & Phelps, 2005; Norrholm, et al., 2006; Van Damme et al., 2006). The current study specifically addresses return of fear as a result of reinstatement.

There are several models of reinstatement. One theory proposes that return of fear to the extinguished CS occurs when the US is reinstated only in the same context in which it is tested (Bouton et al., 2006). This theory states that when the CS+ is tested after reinstatement, there will be increased fear responding to the CS+ only if reinstatement occurred in the same context as the CS+ test context (Bouton, 1984; Bouton & King, 1983; Frohardt et al., 2000, Wilson et al.,

1995). Though excitatory conditioning is independent of context, inhibitory learning is dependent on context (Bouton, 2004). Because the CS+ carries both excitatory and inhibitory meanings after extinction, context helps disambiguate which meaning the CS+ carries in a given moment. For example, an unsignaled US presented in Context A provides an excitatory meaning to Context A, whereas Context B – which did not have an unsignaled US – carries an inhibitory meaning. When the CS+ is tested in Context A, there will be increased fear responding compared to when it is presented in Context B (Bouton, et al., 2006; Bouton, 1984; Bouton & Bolles, 1979; Bouton & King, 1983). Evidence that exposure to the excitatory context reduces the fear-provoking effects of reinstatement further supports this theory (e.g., Bouton and Bolles, 1979). A second theory suggests that an unsignaled US elicits physiological arousal similar to the arousal experienced during fear acquisition and acts as an internal retrieval cue of the excitatory CS/US association (Haroutunian & Riccio, 1979). Context includes exteroceptive (e.g., a room, place, environment, or other external background stimuli; Bouton, 1993) and interoceptive cues, such as drug state (Bouton et al., 1990; Overton, 1985), hormonal state (Ahlers & Richardson, 1985), mood state (Bower, 1981; Eich, 1995), food deprivation state (Davidson, 1992), recent events (Bouton, et al., 1993; Ricker & Bouton, 1996), expectation of events (Bouton et al., 1993), and passage of time (Bouton, 1993; Rosas & Bouton, 1998). This suggests that increased arousal levels after reinstatement that approximate the arousal level of fear acquisition may indeed contribute to reinstatement of fear by acting as an excitatory contextual cue (Haroutunian & Riccio, 1979). Other models exist, as well (e.g., Schmajuk, Larrauri, & LaBar, 2007; Westbrook, et al., 2002). However, none fully explain recent evidence regarding the role of CS+ valence in reinstated fear (Dirikx et al., 2004; Dirikx et al., 2007; Hermans, et al., 2005).

Several studies have shown that the more negatively the CS+ is valenced at the end of extinction, the stronger the conditional fear after reinstatement (i.e., reinstatement test; Dirikx et al., 2004; Dirikx et al., 2007; Hermans, et al., 2005), although the results of one study did not support this association (Dirikx, 2006, Study 3). There is little evidence about the effects of CS+ valence on spontaneous recovery, but the available evidence suggests that CS+ valence does not reliably predict spontaneous recovery (Dirikx et al., 2004; Dirikx et al., 2007; Hermans, et al., 2005). However, whether CS+ valence predicts spontaneous recovery is unclear because these studies did not conduct a well-controlled test of spontaneous recovery. We are also unaware of any studies that investigate whether post-extinction CS+ valence predicts renewal of fear, though one study showed that a positively valenced retrieval cue reduced renewal compared to absence of the positively valenced retrieval cue (Dibbets & Maes, 2011). In this same study, the negatively valenced retrieval cue did not reduce renewal compared to absence of the negatively valenced retrieval cue, and comparing the effects of the positively valenced retrieval cue and the negatively valenced retrieval cue on renewal of fear resulted in no significant differences. Though this does not directly examine whether post-extinction CS+ valence predicts renewal of fear, it suggests that the valence of retrieval cues may affect renewal of fear; any other specific effects of CS+ valence in renewal would require further investigation. Moreover, despite not accounting for the various types of return of fear, similar results to the valence-reinstatement effect have been demonstrated in a clinical sample of individuals with public speaking anxiety. In this study, negative implicit attitudes towards public speaking at the end of exposure therapy predicted return of fear, much like post-extinction CS+ valence predicts reinstated fear (Vasey, et al., 2012). In sum, the association between CS+ valence and return of fear is a relatively new area of interest. Though studies investigating post-extinction CS+ valence and renewal are non-

existent, post-extinction CS+ valence seems to predict reinstatement fear, not spontaneous recovery.

Hermans and colleagues (e.g., Dirikx, et al., 2004) proposed that the network model of emotions (Lang et al., 1990) may account for the valence-reinstatement phenomenon. In this model, emotions are located on a 2 (Valence: positive, negative) x 2 (Arousal: high, low) matrix where valence and arousal are orthogonal. Fear and anxiety belong in the negative valence/high arousal quadrant. Lang, et al. (1993) investigated the covariation between several measures of valence and arousal with the International Affective Picture System (IAPS). As measured by self-report on a 0-29 scale, fear has been shown to have a valence of 7.9 and arousal level of 22.6 (Lang, et al., 1993). Compared to happiness, fear has significantly more negative valence than two subtypes of happiness: erotic (23.3) and nurturant (25.6). Fear's arousal level is significantly greater than that of happy/nurturant (15.9) but not happy/erotic (21.2; see Lang, et al., 1993 for full list of comparisons). Extinction learning decreases arousal towards the CS+, as shown by lower skin conductance response (SCR) – a measure of arousal only, not valence (Bradley, Cuthbert, & Lang, 1990; Cook, Hawk, Davis, & Stevenson, 1991; Greenwald, Cook, & Lang, 1989; Manning & Melchiori, 1974; Winton, Putnam, & Krauss, 1984). However, though CS+ valence may become somewhat more positive from the end of acquisition to the end of extinction, it typically remains more negative than pre-acquisition CS+ ratings and post-extinction CS- valence ratings (Dirikx et al., 2004). Reinstatement may increase arousal and thus return emotions towards the CS+ back into the negative valence/high arousal quadrant (Dirikx, et al., 2004; Dirikx, et al., 2007). The valence-reinstatement model raises the possibility that strategies designed to increase positive valence of the feared CS+ during exposure therapy may ultimately reduce relapse via reinstatement.

However, before a strong case for modifying clinical interventions based on this theory can be made, further experimental work is needed to establish the robustness of the findings and their specificity. The goal of the current study was to evaluate valence of the CS+ at the end of extinction as a predictor of conditional fear following reinstatement. In the present study, CS+ valence is defined as how positive or negative the CS+ is to the individual (see Materials for operational definition). This differs from arousal, which is defined as how exciting or calming the CS+ is to the individual (measured by SCR; see Materials). A second goal was to test specificity of effects by evaluating whether CS+ fear following reinstatement was predicted by post-extinction CS+ fear and/or post-extinction CS+ valence. Specificity was further tested by evaluating whether post-extinction CS+ valence predicted extinction retest CS+ fear as well as at reinstatement test. In accordance with the valence-reinstatement model, we hypothesized that a) post-extinction CS+ valence would be a stronger predictor of reinstatement test CS+ fear than post-extinction CS+ fear, b) post-extinction CS+ valence would predict CS+ fear at reinstatement test but not at extinction retest nor with non-reinstatement control participants.

Methods

Participants

One-hundred and seven undergraduates at the University of California at Los Angeles (UCLA) participated for course research credit or payment of sixty dollars. Data from 22 participants was excluded because they discontinued participation after Day 1 (N = 3) or Day 2 (N = 19), and data from one participant was excluded because of technical difficulties (i.e., the CS+ and CS- were the same image). Thus, the final participant count was 84. Participants were 55% female, and mean age was 19.22 (1.04) years. All but one participant reported ethnicity: Asian (31.32%), Caucasian (42.17%), Hispanic (12.05%), and mixed (14.46%). The study was

approved by UCLA's institutional review board, and all participants were provided with a description of the study and gave written, informed consent.

Materials

Self-Report Measures. An 11-point Likert scale was used to obtain subjective fear ratings of the CS+ and CS− (0 = 'not at all fearful of', 10 = 'very fearful of') and valence ratings of the CS+ and CS- (i.e., 0 = 'not at all unpleasant,' 10 = 'very unpleasant'). These were measured after habituation and acquisition, before and after extinction, before extinction retest, and after test (i.e., either reinstatement test or non-reinstatement control test, depending on randomization). Participants' expectancy of experiencing the US was rated during CS presentations and inter-trial intervals (ITIs) by using a joystick to move an on-screen pointer along an analog scale between the extremes of 0 = 'certain no stimulation' and 10 = 'certain stimulation' with a midpoint of 5 = 'uncertain'. The scale appeared on screen at specific times, prompting participants to make a rating based on their expectancy of experiencing the US in "the next few moments."

Physiological Measures. SCRs to CS onsets served as an index of CS association with the US. SCRs were recorded from two 3mm diameter Ag/AgCl electrodes placed on the distal phalanx of the index and middle fingers of the non-dominant hand. The magnitude of SCRs were calculated as the difference between the trough and apex of the skin conductance level curve, expressed in microsiemens (μS), commencing within 1–4 s following CS onset. SCRs were rejected for a given CS presentation if behavioral observations or other physiological measures indicated excessive drowsiness, movement, or behavior such as coughing and sneezing. This was determined by the experimenter for every CS presentation. SCRs were scored as zero for a given CS presentation when there was no observable SCR activity commencing within the 1–4 s

window. Physiological data was acquired using a Grass Instruments Amplifier System and were digitized and sampled at 1000 Hz.

Apparatus, CSs, and US

CSs, ITIs, US delivery, US expectancy ratings, and recording of physiological activity were under the control of National Instruments LabVIEW Programming Software (v7). CSs consisted of a green triangle or purple trapezoid displayed on a 21-inch computer monitor located 3 feet from participants at eye level. Bicep muscle stimulation, which served as the aversive US, consisted of 20.4mA peak current (equating to a 50V peak) passing between two pads for 0.5 s and was delivered by a Digital 807 Electrical Muscle Stimulation Device (Everyway Medical Instruments). Such stimulation results in a rapid onset, involuntary muscle contraction across the biceps. The intensity level of the stimulation was preset based on pilot testing to a level that was considered uncomfortable but not painful.

Procedures

The study was a 3 (Extinction CS Duration: 1min, 2min, 4min) x 2 (Reinstatement: yes, no) design (see Table 1). Participants were asked to participate in three sessions conducted on three consecutive days with start times differing by no more than 4 hours. Differential fear conditioning was conducted on Day 1. An extinction phase where participants were randomly assigned to one of three groups (1 min, 2min, or 4min CS durations) was conducted on Day 2. This was designed to observe the effects of number of CS+ extinction trials on extinction learning while holding constant the total time exposed to the CS+ (i.e., four CS+ trials of one minute each, two trials of two minutes each, one trial of four minutes; see Prenoveau, et al., 2013). This design feature was not central to the current set of analyses, but was included as a covariate (see Data Analysis section). On Day 3, all participants engaged in extinction retest and

were then were randomized to either reinstatement (in which they experienced an unsignaled US) or the non-reinstatement control condition (which did not include an unsignaled US).

On Day 1, participants underwent habituation and acquisition. The CS+ and CS- were 2-min duration images of a green triangle or purple trapezoid (counterbalanced across participants). During each CS, two online expectancy ratings (early and late) were prompted through presentation of the analog scale at 48 s and 108 s after CS onset. ITIs were 90 s in duration, with the expectancy scale appearing 48 s after CS offset. During acquisition, muscle stimulation (the US) was delivered 117 s after CS+ onset (3 s prior to offset; for full details of Day 1 procedures, see Prenoveau, et al., 2013).

On Day 2, participants underwent extinction. Participants were randomized to one of three groups, which differed on the length of CS presentations as well as number of CS trials (to equate total CS exposure). Participants provided pre-extinction CS fear and valence ratings. The number and duration of CS trials for the three groups was: 1min group: 8, 1min CSs (4 CS+, 4 CS−); 2min group: 4, 2min CSs (2 CS+, 2 CS−); 4min group: 2, 4min CSs (1 CS+, 1 CS−). The number of expectancy ratings per time of CS display was held constant across groups (i.e., four ratings). ITI duration for all three groups was 10 min with three expectancy ratings during each ITI (for full details of Day 2 procedures, see Prenoveau, et al., 2013).

Extinction retest and reinstatement (with a non-reinstatement control condition) were conducted on Day 3. All participants received two 2-min presentations each of the CS+ and CS- in randomized order, a 4-min ITI, and another two 2-min presentations each of the CS+ and CS- in randomized order. Participants were randomized to either receive an unsignaled US during the 4-min ITI (i.e., reinstatement group) or not receive one (i.e., non-reinstatement control group). The timing and number of expectancy prompts for CSs and ITIs were identical to that of

acquisition for both conditions, and self-report fear and valence ratings were acquired before extinction re-test and after the final reinstatement test CS presentation.

Data Analysis

Hierarchical linear regressions were conducted to test the study hypotheses. Because the three extinction conditions that varied by CS duration were not the focus of the present analyses, we controlled for extinction condition in the first regression block in the extinction retest and reinstatement analyses. The second block controlled for 'CS+' fear or 'CS+ minus CS-' fear at the previous time point (i.e., end of extinction for extinction retest analyses; beginning of extinction retest for reinstatement analyses) using the same dependent variable (e.g., SCR at extinction retest when measuring SCR at reinstatement). The third block included either 'CS+' valence or 'CS+ minus CS-' valence at the end of extinction. Dependent measures were calculated congruently with valence ratings: 'CS+' valence was tested as a predictor of 'CS+' fear, and 'CS+ minus CS-' valence was tested as a predictor of 'CS+ minus CS-' fear.

Because of the study design (i.e., varying number of CS+ presentations during extinction), it was not possible to uniformly control for SCR across groups at the end of extinction. Thus, SCR was not analyzed as a dependent variable in the extinction retest analyses. CS valence at the end of extinction was tested as a predictor of US expectancy and self-report CS fear at extinction retest. CS valence at the end of extinction was also tested as a predictor of US expectancy, self-report fear, and SCR at reinstatement test controlling for extinction retest. We also evaluated CS fear at the end of extinction as a predictor of extinction retest fear and reinstatement fear using the same measures as indicated above.

Results

General Fear Conditioning Results

Figure 1 shows CS+ and CS- self-report fear collapsed across extinction groups with dependent samples t-tests comparing CS+ self-report fear to CS- self-report fear within each fear conditioning phase. Using the dependent samples t-tests, there were no significant differences in CS fear at post-habituation ($t(83) = -.440$, $p = .661$), but there was significantly greater fear of the CS+ than the CS- at post-acquisition, as expected ($t(83) = 7.088$, $p < .001$). There was also significantly more fear of the CS+ than the CS- at pre-extinction ($t(82) = 5.285$, $p < .001$), post-extinction ($t(83) = 2.724$, $p = .008$), extinction retest ($t(82) = 4.115$, $p < .001$), and non-reinstatement test for control participants ($t(45) = 2.660$, $p = .011$). However, there was no significant difference in CS+ and CS- self-report fear at reinstatement test for reinstatement participants ($t(39) = 1.071$, $p = .291$). There was a significant decrease in CS+ fear relative to CS- fear in the 1-min extinction group, whereas this was not the case in the 2-min and 4-min groups. There were no differences in CS fear between extinction groups at extinction retest (see Prenoveau, et al., 2013 for details).

To analyze the effects of the presence or absence of reinstatement on CS fear, 2 (Group: Reinstatement, Non-Reinstatement Control) x 2 (CS Type: CS+, CS-) x 2 (Phase: Extinction Retest, Reinstatement Test) repeated measures ANOVAs were conducted with self-report fear, SCR, and US expectancy as the dependent variables. None of the three-way interactions were significant: self-report fear ($F(1, 81 = 2.021$, $p = .159$), SCR ($F(1, 60) = .023$, $p = .879$), and US expectancy ($F(1, 77) = 1.445$, $p = .233$). However, using self-report CS+ fear as the dependent variable and excluding CS type from the analyses, there was a significant 2 (Group: Reinstatement, Non-Reinstatement Control) x 2 (Phase: Extinction Retest, Reinstatement Test) interaction ($F(1, 81) = 7.161$, $p = .009$). Simple effects tests revealed that, at extinction retest, there was no significant difference in CS+ fear between the reinstatement and non-reinstatement

control groups ($F(1, 81) = 1.627$, $p = .206$); similarly, at reinstatement test, there was also no significant difference between the reinstatement and non-reinstatement control groups ($F(1, 81) = 1.100$, $p = .297$). However, non-reinstatement control participants significantly decreased in self-report CS+ fear from extinction retest to non-reinstatement control test ($F(1, 81) = 9.803$, $p = .002$), but there was no significant change in self-report CS+ fear for reinstatement participants from extinction retest to reinstatement test ($F(1, 81) = .573$, $p = .451$).

Using SCR as the dependent variable, there was no significant 2 (Group: Reinstatement, Non-Reinstatement Control) x 2 (Phase: Extinction Retest, Reinstatement Test) interaction ($F(1, 69) = .262$, $p = .610$). Lastly, using US expectancy as the dependent variable, there was no significant 2 (Group: Reinstatement, Non-Reinstatement Control) x 2 (Phase: Extinction Retest, Reinstatement Test) interaction ($F(1, 81) = 2.719$, $p = .103$).

Primary Analyses – Extinction Retest

CS+ Valence as Predictor

Neither 'CS+' nor 'CS+ minus CS-' valence ratings predicted any measure of CS fear at extinction retest (see Tables 2 and 3 for a summary of results and Table 4 descriptive statistics of the measures).

CS+ Fear as Predictor

Higher self-reported 'CS+' fear after extinction predicted higher self-reported 'CS+' fear at extinction retest while controlling for post-extinction 'CS+' valence ($\beta = .608$, $\Delta R^2 = .081$, $t(74) = 3.400$, $p = .001$). Higher self-reported fear of 'CS+ minus CS-' after extinction predicted higher self-reported fear of 'CS+ minus CS-' at extinction retest while controlling for 'CS+ minus CS-' valence after extinction ($\beta = .1.049$, $\Delta R^2 = .141$, $t(74) = 6.163$, $p < .001$).

Higher ‘CS+ minus CS-’ US expectancy after extinction approached significance as a predictor of ‘CS+ minus CS-’ US expectancy at extinction retest when controlling for ‘CS+ minus CS-’ valence after extinction ($\beta = .244$, $\Delta R^2 = .050$, $t(71) = 1.987$, $p = .051$). No other effects were significant.

Primary Analyses – Reinstatement Test

CS+ Valence as Predictor – Reinstatement Condition

More negative valence of the ‘CS+’ at the end of extinction significantly predicted fear of ‘CS+’ at reinstatement test, measured using US expectancy ($\beta = .463$, $\Delta R^2 = .207$, $t(32) = 3.045$, $p = .005$) and self-reported fear ($\beta = .835$, $\Delta R^2 = .393$, $t(32) = 7.209$, $p < .001$) while controlling for fear of ‘CS+’ at the end of extinction. Similarly, more negative valence of ‘CS+ minus CS-’ at the end of extinction significantly predicted higher fear of ‘CS+ minus CS-’ at reinstatement test, measured using US expectancy ($\beta = .491$, $\Delta R^2 = .222$, $t(31) = 3.371$, $p = .002$) and self-reported fear ($\beta = .637$, $\Delta R^2 = .082$, $t(32) = 3.345$, $p = .002$) while controlling for fear of ‘CS+ minus CS-’ at the end of extinction. No other effects were significant.

CS+ Fear as Predictor – Reinstatement Condition

Self-reported fear of ‘CS+’ after extinction approached significance as a predictor of fear of ‘CS+’ at reinstatement test ($\beta = .575$, $\Delta R^2 = .028$, $t(31) = 2.026$, $p = .051$) while controlling for valence of ‘CS+’ at the end of extinction. No other effects were significant.

CS+ Valence as Predictor – Non-Reinstatement Condition

CS valence (using both the ‘CS+’ and ‘CS+ minus CS-’ calculations) after extinction did not significantly predict fear at reinstatement test for the non-reinstatement control condition while controlling for fear of CS at the end of extinction.

CS+ Fear as Predictor –Non-Reinstatement Condition

'CS+' US expectancy after extinction significantly predicted 'CS+' US expectancy at reinstatement test for the non-reinstatement condition ($\beta = .424$, $\Delta R^2 = .122$, $t(33) = 3.086$, $p = .004$) while controlling for valence of 'CS+' at the end of extinction. Self-reported fear of 'CS+ minus CS-' approached significance as a predictor of self-reported 'CS+ minus CS-' fear while controlling for valence of 'CS+ minus CS-' at the end of extinction, though this was an inverse relationship ($\beta = -.608$, $\Delta R^2 = .024$, $t(36) = -1.692$, $p = .099$). No other effects were significant.

Discussion

The purpose of the present study was to investigate whether post-extinction CS+ valence predicted CS+ fear at reinstatement. The tests of reinstatement showed that the non-reinstatement control participants decreased CS+ fear from extinction retest to reinstatement test, whereas the reinstatement participants did not. Also, the reinstatement participants showed non-differential fear of the CS+ and CS- after reinstatement, whereas the non-reinstatement control participants showed greater fear of the CS+ than CS- at reinstatement test. As hypothesized, less negative post-extinction CS+ valence predicted less reinstatement test CS+ fear over and above post-extinction CS+ fear. This finding was robust across US expectancy and self-report fear as measures, as well as across indices of absolute 'CS+' fear and differential 'CS+ minus CS-' fear. Post-extinction CS+ valence did not predict CS+ fear at reinstatement test in the non-reinstatement control condition nor at extinction retest, which suggests that CS+ valence is a specific predictor of fear after reinstatement, not spontaneous recovery. Furthermore, post-extinction CS+ fear was a predictor of CS+ fear (over and above CS+ valence) measured via US expectancy in both the non-reinstatement condition at reinstatement test and measured via self-report at extinction retest. This suggests that post-extinction CS+ valence is a specific predictor of CS+ fear after reinstatement, whereas post-extinction CS+ fear predicts CS+ fear at extinction retest (i.e., spontaneous recovery).

The current findings partly support the valence-reinstatement theory (Dirikx, et al., 2004; Dirikx, et al., 2007; Hermans, et al., 2005). The theory states that return of fear via reinstatement occurs from a combination of residual negative CS+ valence after extinction and an increase in arousal resulting from the unsignaled US. The current study focused on the "valence" aspect of the valence-reinstatement theory but did not address the arousal aspect. Because of the present study's focus on valence, we were also not able to fully evaluate the valence-reinstatement theory in comparison with the context-specific model of reinstatement (e.g., Bouton, et al., 2006) nor the arousal model (Haroutunian & Riccio, 1979). Future studies could measure increases in arousal specifically to the context after reinstatement as well as arousal specifically to the CS+. Similarly, future studies would benefit from further testing the effects of implicit (e.g., postauricular reflex and eye blink startle reflex to startle probe; Benning, et al., 2004; Sandt, et al., 2009; Personalized Implicit Association Test; Vasey, et al., 2012) and explicit measures of CS+ valence (e.g., self-report) in the prediction of reinstatement CS+ fear.

The present study also has relevant clinical implications. Because return of fear is a major challenge currently faced by clinicians in the treatment of anxiety disorders, ranging from 19 to 62% of cases (Craske & Mystkowski, 2006), understanding the mechanisms of return of fear is critical. The present study suggests that more negative post-extinction CS+ valence – which parallels CS+ valence at the end of treatment or end of a therapy session – predicts more return of fear via reinstatement. Efforts focused on increasing CS+ valence in therapy sessions may help reduce relapse via reinstatement. However, before this should be implement in treatment, further research replicating and expanding on the present study is needed. To this end, future experimental studies could investigate whether methods of increasing positive post-extinction

CS+ valence reduce reinstatement fear compared to control. For example, positive imagery training (Holmes, et al., 2006) has reliably been demonstrated to increase positive mood. Positive mood induction has been shown to: a) increase positive valence toward a specific stimulus (Erez, et al., 2002), b) increase recall of positive stimuli (Teasdale, et al., 1983), and c) decrease time to retrieve positive memories (Teasdale, et al., 1979). Thus, conducting positive imagery training and increasing positive mood before extinction could increase positive valence toward the CS+ and reduce reinstated fear. Alternatively, attempts to directly increase positive valence towards the CS+ may also show efficacy (e.g., in the case of spider phobia, developing methods of training clients to like spiders more).

Moreover, little is known of the effect of introducing a novel US at reinstatement on CS+ fear. In one animal study, a novel US at reinstatement (i.e., a klaxon) different from the original CS/US pairing (i.e., electric shock) reinstated fear of the CS+ in rats (Rescorla & Heth, 1975). Similar results exist with humans when using an electric shock and loud noise as the USs (Sokol & Lovibond, 2012). In the latter study, both the original and novel USs increased SCR to the CS+ at reinstatement test, but there was only an increase in US expectancy to the novel US, not the original US. If, for example, an individual with social anxiety disorder fears giving public speeches (the CS+) for fear of experiencing negative social evaluation (the US), this may mean that experiencing reinstatement with a different US (e.g., physical pain/bodily harm from a car accident) may increase fear of giving public speeches. Future studies on this topic would both help elucidate the effect of a novel US on CS+ fear at reinstatement, as well as whether post-extinction CS+ valence predicts reinstated CS+ fear after presentation of a novel US. Investigating this issue could prove clinically useful for treating clients who are faced with aversive events that are not of the same nature as the US they associate with the CS+.

There were several limitations of the present study. First, the study design included three different extinction methods (Group 1: four 1-min CS+s; Group 2: two 2-min CS+s; Group 3: one 4-min CS+), the effects of which were controlled statistically. Second, we were not able to predict return of fear at extinction retest using SCR. Third, aversive shocks as the US were set at a pre-determined intensity level based on pilot data rather than conducting a work-up procedure to reach a pre-determined subjective rating of discomfort (e.g., 7 out of 10 signifying "uncomfortable but not painful"). Uniform intensity level likely induced variability in the subjective level of discomfort experienced by each participant and thus may have affected CS+ fear and valence levels. Fourth, the self-report valence scale ranged from 0 = "not at all unpleasant" to 10 = "very unpleasant." A value of "not at all unpleasant" does not necessarily mean that there was positive valence associated with the CS but rather the absence of unpleasantness (i.e., absence of negative valence). Thus, future studies could improve upon this by including a valence scale that ranges from "very unpleasant" to "very pleasant." Lastly, there was only one US presentation at reinstatement, whereas other studies have used two (Dirikx, et al., 2004; Dirikx, et al., 2007) or four (Hermans, et al., 2005) US presentations. Related is that the significant 2 (Group: Reinstatement, Non-Reinstatement Control) x 2 (Phase: Extinction Retest, Reinstatement Test) interaction was driven by self-report CS+ fear reduction in the non-reinstatement control group rather than an increase in self-report CS+ fear in the reinstatement group. Stronger reinstatement effects may have been obtained with more US presentations.

In conclusion, the present study suggests that CS+ valence at the end of extinction is a reliable and robust predictor of CS+ fear at reinstatement. Vasey et al. (2012) demonstrated this effect in a clinical setting, which has implications for predicting and preventing relapse after treatment. Future studies would benefit both the basic scientific fear conditioning literature and

clinical anxiety literature by further elucidating this relationship and investigating methods of increasing the positive valence of the CS+ at the end of extinction.

References

1. Ahlers ST, Richardson R. 1985. Administration of dexamethasone prior to training blocks ACTH-induced recovery of an extinguished avoidance response. *Behav Neurosci* **99:** 760 – 764.
2. Benning SD, Patrick CJ, Lang AR. 2004. Emotional modulation of the post-auricular reflex. *Psychophysiology* **41:** 426-432.
3. Bouton ME. 2004. Context and behavioral processes in extinction. *Learn Mem* **11:** 485-494.
4. Bouton ME. 1993. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning, *Psychol Bull* **114:** 80-99.
5. Bouton ME. 2002. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction, *Biol Psychiatry* **52:** 976-986.
6. Bouton ME. 1984. Differential control by context in the inflation and reinstatement paradigms. *J Exp Psychol Anim Behav Process* **10:** 56-74.
7. Bouton ME, Bolles, RC. 1979. Contextual control of the extinction of conditioned fear. *Learn Motiv*, **10**: 445-466.
8. Bouton ME, Kenney FA, Rosengard C. 1990. State-dependent fear extinction with two benzodiazepine tranquilizers. *Behav Neurosci* **104:** 44 –55.
9. Bouton ME, King DA. 1983. Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *J Exp Psychol: Anim Behav Process* **9:** 248–65.
10. Bouton ME, Rosengard C, Achenbach GG, Peck CA, Brooks DC. 1993. Effects of contextual conditioning and unconditional stimulus presentation on performance in appetitive conditioning. *Q J Exp Psychol B* **46:** 63–95.
11. Bouton ME, Westbrook RF, Corcoran KA, Maren S. 2006. Contextual and temporal modulation of extinction: behavioral and biological mechanisms. *Biol Psychiatry*, **60:** 352-360.
12. Bower GH 1981. Mood and memory. *Am Psychol* **36:** 129 –148.
13. Bradley MM, Cuthbert BN, Lang PJ. 1990. Startle reflex modification: Emotion or attention? *Psychophvsiology*, **27:** 513-52
14. Campbell DT, Fiske DW. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* **56.2:** 81.
15. Cook EW III, Hawk LH, Davis TL, Stevenson VE. 1991. Affective individual differences and startle reflex modulation. *J Abnormal Psychol* **100:** 5-13.
16. Craske MG, Mystkowski JL. 2006. Exposure therapy and extinction: clinical studies. In *Fear and Learning: From Basic Processes to Clinical Implications*, (ed. MG Craske, D Hermans, D Vansteenwegen), pp. 217–33. Washington, DC: Am. Psychiatr. Assoc.
17. Culver NC, Stoyanova M, Craske MG. 2011. Clinical relevance of retrieval cues for attenuating context renewal of fear. *J Anxiety Disord* **25:** 284-292.
18. Davidson TL, Flynn FW, Jarrard, LE. 1992. Potency of food deprivation intensity cues as discriminative stimuli. *J Exp Psychol: Anim Behav Process* **18:** 174-181.

19. Dibbets P, Maes, JH. 2011. The effect of an extinction cue on ABA-renewal: Does valence matter? *Learn Motiv* **42*:*** 133-144.
20. Eich E. 1995. Mood as a mediator of place dependent memory. *J Exp Psychol Gen* **124:** 293–308.
21. Erez A, Isen AM. 2002. The influence of positive affect on the components of expectancy motivation. *J Appl Psychol* **87:** 1055-1067.
22. Frohardt RJ, Guarraci FA, Bouton ME. 2000. The effects of neurotoxic hippocampal lesions on two effects of context after fear extinction. *Behav Neurosci* **114:** 227.
23. Greenwald MK, Cook EW III, Lang PJ. 1989. Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *J Psychophysiology* **3:** 51-64.
24. Haroutunian V, Riccio DC. 1979. Drug-induced "arousal" and the effectiveness of CS exposure in the reinstatement of memory. *Behav Neur Biol*, **26*:*** 115-120.
25. Hofmann SG, Smits JAJ. 2008. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials, *J Clin Psychiatry* **69:** 621-632.
26. Holmes EA, Mathews A, Dalgleish T, Mackintosh B. 2006. Positive interpretation training: Effects of mental imagery versus verbal training on positive mood. *Behav Ther* **37:** 237-247.
27. In-Albon T, Schneider S. 2007. Psychotherapy of childhood anxiety disorders: a meta-analysis, *Psychother Psychosom* **78:** 15-24.
28. LaBar K, Phelps EA. 2005. Reinstatement of conditioned fear in humans is context dependent and impaired in amnesia. *Behav Neurosci.* **119:** 677–86.
29. Lang PJ, Bradley MM, Cuthbert BN. 1990. Emotion, attention, and the startle reflex. *Psychol Bull* **97:** 377–95
30. Lang PJ, Greenwald MK, Bradley MM, Hamm AO. 1993. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, **30:** 261-273.
31. Manning S., Melchiori MP. 1974. Words that upset urban college students: Measured with GSRs and rating scales. *J Soc Psychol* **94:** 305-306.
32. Myers KM, Davis M. 2007. Mechanisms of fear extinction. *Mol Psychiatry* **12:** 120-150.
33. Norrholm SD, Jovanovic T, Vervliet B, Myers K, Davis M, Rothbaum BO, Duncan EJ. 2006. Conditioned fear extinction and reinstatement in a human fear potentiated startle paradigm. *Learn Mem* **13:** 681–85.
34. Overton DA. 1985. Contextual stimulus effects of drugs and internal states. In *Context and Learning* (ed. PD Balsam, A Tomie), pp. 357-84. Hillsdale, NJ: Erlbaum.
35. Prenoveau JM., Craske MG., Liao B, Ornitz EM. 2013. Human fear conditioning and extinction: Timing is everything… or is it?. *Biol Psychol* **92:** 59-68.
36. Quirk GJ. 2002. Memory for extinction of conditioned fear is long-lasting and persists following spontaneous recovery. *Learn Mem* **9:** 402–7.
37. Rachman S. 1989. The return of fear: Review and prospect. *Clin Psychol Rev* **9:** 147-168.
38. Rescorla RA, Heth D. 1975. Reinstatement of fear to an extinguished conditioned stimulus. *J Exp Psychol: Anim Behav Process* **104:** 88–96.

39. Ricker ST, Bouton ME. 1996. Reacquisition following extinction in appetitive conditioning, *Anim Learn Behav* **24:** 423-436.
40. Rosas JM, Bouton ME. 1998. Context change and retention interval can have addititve, rather than interactive, effects after taste aversion extinction, *Psychon Bull Rev* **5:** 79-83.
41. Sandt AR, Sloan DM, Johnson KJ. Measuring appetitive responding with the postauricular reflex. *Psychophysiology*, **46:** 491-497.
42. Schmajuk NA, Larrauri JA, LaBar KS. 2007. Reinstatement of conditioned fear and the hippocampus: An attentional-associative model. *Behav Brain Res* **177:** 242-253.
43. Sokol N, Lovibond PF. 2012. Cross-US reinstatement of human conditioned fear: return of old fears or emergence of new ones? *Behav Res Ther* **50:** 313–22.
44. Teasdale JD, Fogarty SJ. 1979. Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *J Abnormal Psychol* **88:** 248-257.
45. Teasdale JD, Russell ML. 1983. Differential effects of induced mood on the recall of positive, negative and neutral words. *Br J Clin Psychol* **22:** 163-171.
46. Van Damme S, Crombez G, Hermans D, Koster EHW, Eccleston C. 2006. The role of extinction and reinstatement in attentional bias to threat: a conditioning approach. *Behav Res Ther* **44:** 1555–63.
47. Vasey MW, Harbaugh CN, Buffington AG, Jones CR, Fazio RH. 2012. Predicting return of fear following exposure therapy with an implicit measure of attitudes. *Behav Res Ther* **50:** 767-774.
48. Wagner AR. 1981. SOP: A model of automatic memory processing in animal behavior. In *Information Processing in Animals: Memory Mechanisms.* (ed. NE Spear, RR Miller), pp 5–47. Erlbaum, Hillsdale, NJ.
49. Westbrook RF, Iordanova M, McNally G, Richardson R, Harris JA. 2002. Reinstatement of fear to an extinguished conditioned stimulus: two roles for context. *J Exp Psychol Anim Behav Process* **28:** 97.
50. Wilson A, Brooks DC, Bouton ME. 1995. The role of the rat hippocampal system in several effects of context in extinction. *Behav Neurosci* **109:** 828.
51. Winton WM, Putnam LE, Krauss RM. 1984. Facial and autonomic manifestations of the dimensional structure of emotion. *J Exp Soc Psychol* **20:** 195-216.

Table 1. Overview of the experimental procedure with order of tasks listed from top to bottom within Day.

Day 1	Day 2			Day 3	
All Participants	**1-min Group**	**2-min Group**	**4-min Group**	**Reinstatement Group**	**Non-Reinstatement Control Group**
Habituation	Extinction	Pre- or Post-Extinction Period	Pre- or Post-Extinction Period	Extinction Retest	Extinction Retest
2 'to-be' CS+ (2 min each)	Valence and Fear Ratings	Valence and Fear Ratings	Valence and Fear Ratings	Valence and Fear Ratings	Valence and Fear Ratings
2 'to-be' CS- (2 min each)	4 CS+ (1 min each)	40 min (4, 10-min 'ITI-like' blocks)	60 min (6, 10-min 'ITI-like' blocks)	2 CS+ (2 min each)	2 CS+ (2 min each)
3 ITIs (1.5 min each)	4 CS- (1 min each)	Extinction	Extinction	2 CS- (2 min each)	2 CS- (2 min each)
Valence and Fear Ratings	7 ITIs (10 min each)	2 CS+ (2 min each)	1 CS+ (4 min each)	Reinstatement	Non-Reinstatement Control
Acquisition	Valence and Fear Ratings	2 CS- (2 min each)	1 CS- (4 min each)	4-min ITI with one unsignaled US	4-min ITI without US
4 CS+ (2 min each)		3 ITIs (10 min each)	3 ITIs (10 min each)	Reinstatement Test	Non-Reinstatement Control Test
4 CS- (2 min each)		Valence and Fear Ratings	Valence and Fear Ratings	2 CS+ (2 min each)	2 CS+ (2 min each)
7 ITIs (1.5 min each)				2 CS- (2 min each)	2 CS- (2 min each)
Valence and Fear Ratings				Valence and Fear Ratings	Valence and Fear Ratings

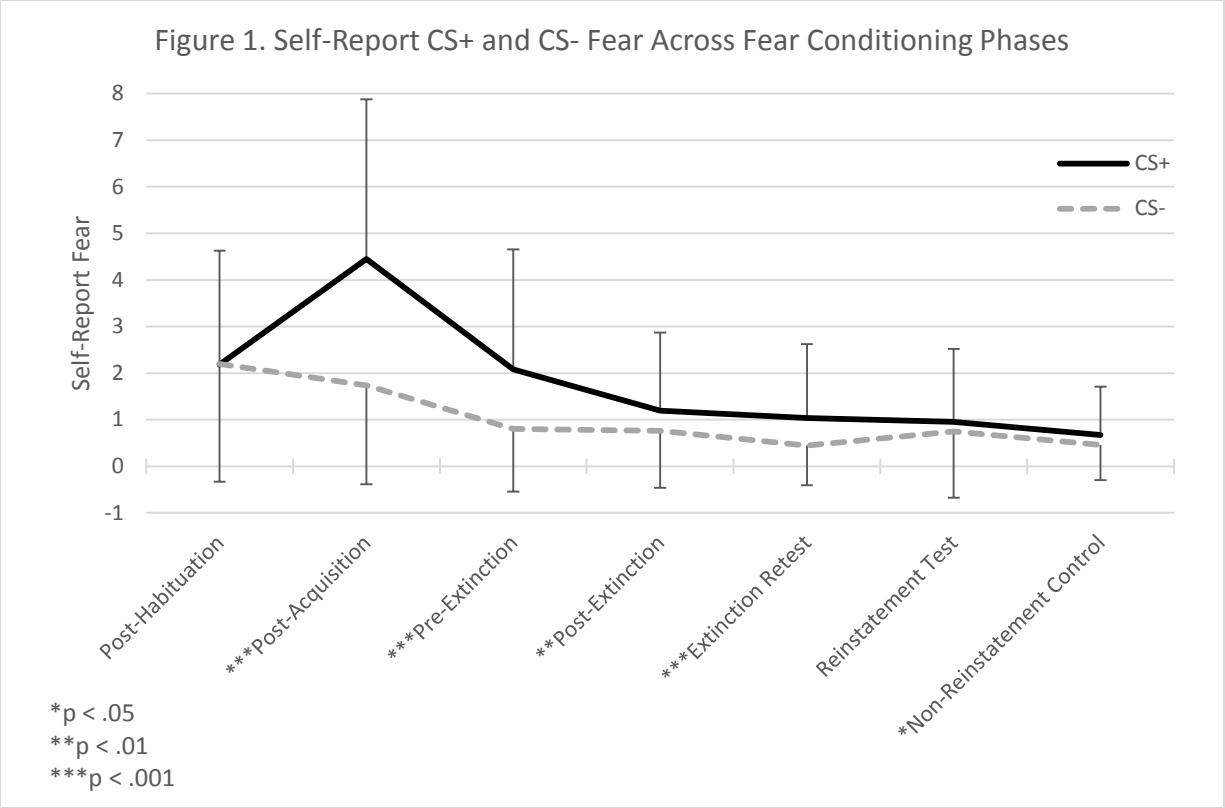
Figure 1. Self-Report CS+ and CS- Fear Across Fear Conditioning Phases
Self-Report Fear
8
7
6
5
4
3
2
1
0
-1
CS+
CS-
Post-Habituation
***Post-Acquisition
***Pre-Extinction
**Post-Extinction
***Extinction Retest
Reinstatement Test
*Non-Reinstatement Control
*p < .05
**p < .01
***p < .001

Table 2. P-Values of Post-Extinction CS Valence Predicting Extinction Retest and Reinstatement Test CS Fear While Controlling for Post-Extinction Fear

		'CS+'	'CS+ minus CS-'
Extinction Retest	US Expectancy	0.137	0.311
	Self-Report Fear	0.592	0.199
Reinstatement Test	SCR	0.693	0.631
	US Expectancy	0.005*	0.002*
	Self-Report Fear	<.001*	0.002*
Non-Reinstatement Control	SCR	0.144	0.626
	US Expectancy	0.744	0.540
	Self-Report Fear	0.518	0.923

* = Significant

Table 3. P-Values of Post-Extinction CS Fear Predicting Extinction Retest and Reinstatement Test CS Fear Controlling for Post-Extinction CS Valence

		'CS+'	'CS+ minus CS-'
Extinction Retest	US Expectancy	0.245	0.051
	Self-Report Fear	0.001*	<.001*
Reinstatement Test	SCR	-	-
	US Expectancy	0.263	0.102
	Self-Report Fear	0.051	0.809
Non-Reinstatement Control	SCR	-	-
	US Expectancy	0.004*	0.514
	Self-Report Fear	0.519	0.099

***** = Significant

Table 4. Independent Variable and Dependent Variable Means and Standard Deviations

	Post-Habituation	Post-Acquisition	Pre-Extinction	Post-Extinction	Extinction Retest	Reinstatement Test	Non-Reinstatement Control
				CS+			
SCR	.017 (.033)	.025 (.043)	.058 (.066)	-	.047 (.064)	.028 (.052)	.036 (.053)
US Expectancy	2.157 (8.002)	5.171 (8.408)	5.756 (7.778)	5.454 (7.977)	4.854 (7.741)	3.193 (7.503)	3.181 (7.853)
Self-Report Fear	2.181 (2.449)	4.452 (3.427)	2.081 (2.577)	1.195 (1.676)	1.035 (1.590)	.95 (1.57)	.674 (1.034)
Self-Report Valence	2.600 (2.673)	4.885 (3.414)	2.337 (2.721)	1.414 (1.709)	1.174 (1.696)	1.13 (1.59)	1.000 (1.300)
				CS+ - CS-			
SCR	0 (.027)	.011 (.049)	0 (.062)	-	0 (.077)	.01 (.07)	.020 (.051)
US Expectancy	.153 (1.809)	1.259 (2.798)	1.043 (3.326)	1.510 (2.566)	1.198 (2.872)	.029 (1.953)	.782 (1.935)
Self-Report Fear	-.019 (1.168)	2.144 (2.704)	1.280 (2.146)	.440 (1.420)	.590 (1.314)	.2 (1.81)	.220 (.554)
Self-Report Valence	.019 (1.028)	2.260 (2.745)	1.350 (2.294)	.379 (1.349)	.510 (1.308)	.18 (1.38)	.280 (.779)